

EDITORIAL

Open Access

Hardware implementation of machine vision systems: image and video processing

Guillermo Botella^{1*}, Carlos García¹ and Uwe Meyer-Bäse²

Abstract

This contribution focuses on different topics covered by the special issue titled 'Hardware Implementation of Machine vision Systems' including FPGAs, GPUS, embedded systems, multicore implementations for image analysis such as edge detection, segmentation, pattern recognition and object recognition/interpretation, image enhancement/restoration, image/video compression, image similarity and retrieval, satellite image processing, medical image processing, motion estimation, neuromorphic and bioinspired vision systems, video processing, image formation and physics based vision, 3D processing/coding, scene understanding, and multimedia.

Keywords: Computer vision; Hardware implementation; Image processing; Video processing

Introduction

Machine Vision systems represent an extensive area belonging to the digital signal processing with vast applications nowadays for industrial environment, robotics, vehicular/avionics/space technology, multimedia, entertainment, security, medicine, scientific, and so on. Nevertheless, mostly these techniques cannot be achieved "off-line" due to the physical constraints of the complex environment and the context where they are applied. Additionally, as these systems are becoming complex, huge computational resources are demanding, mostly if hard real-time constraint is required.

The hardware implementation and the acceleration of these systems are gaining importance nowadays in order to accomplish with the requirements of the systems being usually needed to keep a scrupulous trade-off between accuracy, efficiency, and power consumption. The use of GPGPUs, cells, FPGAs, and DSPs as accelerators, thus, achieves speedups of orders of magnitude vs. any optimized CPU implementation. Each appropriate implementation solution is carried out depending on many final demands as power consumption, performance, accuracy, reliability, rapid prototyping, final cost, capability, and reconfigurability among others [1-6].

The objective of this special issue has been to bring together high-quality state-of-the-art research contributions as well as review articles that address recent developments in the hardware implementation of machine vision systems.

Specific contributions

In this special issue, we have detailed fundamental aspects related to hardware acceleration of image and video processing through different heterogeneous systems as commented previously. The applications described in this issue covers from multimedia coding to many machine vision topics.

Regarding this last aspect [7,8], we can remark that low-level vision gets useful measurements such as color, spatial frequency, binocular disparity, motion processing, etc., from several channels. Some of the aforementioned channels or space-temporal filters can be identified with receptive fields that deliver information to the retina. Others, such as binocular disparity or motion processing, are combinations of the previously mentioned ones. Mid-level vision integrates primitives processes at a previous level. Information delivered at this stage corresponds to real-world inferences such as egomotion and independent moving objects (IMOs). They are called causal actions or object candidates in connection with any multimodal characterization. Examples of these are the combination of luminance measurements to infer lightness, shape from shading, perceptual

* Correspondence: gbotella@fdi.ucm.es

¹Department of Computer Architecture and Automation, Complutense University, Ciudad Universitaria s/n, 28040 Madrid, Spain

Full list of author information is available at the end of the article

grouping, figure organization, etc. Finally, high-level vision interprets the scene through specific tasks such as relational reasoning, knowledge building, object recognition, etc.

As resumed, in the initial classification, two principal topics are covered in 12 top quality research articles:

1. New advances in hardware acceleration techniques for multimedia content representation/coding and transmission
2. Hardware acceleration techniques for machine vision systems belonging to low-, mid- and high-level vision (according to the aforementioned taxonomy)

The first topic is covered by the five following contributions:

- In the article entitled 'MPCM: a hardware coder for super slow motion video sequences' [9] by Estefania Alcocer et al., the authors develop a fast FPGA implementation of a simple codec called modulo-pulse code modulation (MPCM) which is able to reduce the bandwidth requirements up to 1.7 times at the same image quality when compared with PCM coding. This allows current high-speed cameras to capture in a continuous manner through a 40-Gbit Ethernet point-to-point access.
- In the article entitled 'Enhancing LTW image encoder with perceptual coding and GPU-optimized 2D-DWT transform' [10] by Miguel Martínez-Rach et al., the authors propose an optimization of the wavelet image coder (specifically E_LTW encoder) with the aim to increase its rate/distortion (R/D) performance through perceptual encoding techniques and reduce the encoding time by means of a graphics processing unit-optimized version of the two-dimensional discrete wavelet transform. The results show that in both performance dimensions, our enhanced encoder achieves good results compared with Kakadu and SPIHT encoders (optimized versions of JPEG2000), achieving speedups of six times.
- In the article entitled 'GPU-based 3D lower tree wavelet video encoder' [11] by Vicente Galiano et al., the authors introduce a fast GPU-based encoder which uses 3D-discrete wavelet transform and lower trees, presenting also an exhaustive analysis of the use of GPU memory. This approach demonstrates a good trade-off between R/D, coding delay (as fast as MPEG-2 for high definition) and memory requirements (up to six times less memory than $\times 264$ encoder).
- In the article entitled 'Adapting hierarchical bidirectional inter prediction on a GPU-based

platform for 2D and 3D H.264 video coding' [12] by Rafael Rodríguez-Sánchez et al., the authors present an algorithm that concurrently performs the inter prediction carried out over *P*- and *B*-frames. The approach implements the hierarchical B frame prediction implemented in the H.264/AVC JM 17.2 reference software encoder, and it is tested using the main and the stereo high profile. Thus, a GPU-based implementation of an H.264/AVC and H.264/MVC inter prediction algorithm on a graphics processing unit is developed. The results show a negligible rate distortion drop with a time reduction of up to 98% for the complete H.264/AVC encoder.

- The last article regarding this first topic is entitled 'Modeling of a method of parallel hierarchical transformation for fast recognition of dynamic images' [13] by Leonid Timchenko et al.; the authors present principles necessary to develop a method, and computational facilities for the parallel hierarchical transformation based on high-performance GPUs are discussed in the paper. Mathematical models of the parallel hierarchical (PH) network training for the transformation and a PH network training method for recognition of dynamic images are developed.

The second topic is discussed by the seven following contributions:

- Filter design optimization is a low-level vision task that has been covered in the paper entitled 'FIR Filter optimization for video processing on FPGAs' [14] by Michael Kumm et al.; the authors develop two proposed optimization techniques for high-speed implementations of the required multiplications in FIR filters with the least possible number of FPGA components. Both methods use integer linear programming formulations which can be optimally solved by standard solvers. In the first method, a formulation for the pipelined multiple constant multiplication problem is presented. In the second method, multiplication structures based on look-up tables are also taken into account. Due to the low coefficient word size in video processing, filters of typically 8 to 12 bits of an optimal solution is found for most of the filters in the benchmark used. A complexity reduction of 8.5% is achieved compared to the state-of-the-art heuristics.
- Super-resolution is also a low-level vision task that has been detailed in the paper entitled 'Exploring super-resolution implementations across multiple platforms' [15] by Bryan Leung et al. The authors implement and analyze a super-resolution algorithm across multiple platforms ranging from purely

hardware to purely software and even a mix of both hardware and software. More specifically, they examine the performance for a FPGA implementation on two different FPGAs, a software/hardware solution on a FPGA with a soft core processor (embedded system), a GPGPU implementation, and a MATLAB implementation. Overall, they are very interesting in terms of architecture and performance over different platforms.

- Regarding low-level vision, the optimization of hyperspectral imaging is a relevant topic nowadays belonging to remote sensing, that has been addressed in the paper entitled 'Performance versus energy consumption of hyperspectral unmixing algorithms on multi-core platforms' [16] by Alfredo Remón et al. The authors provide a detailed assessment of the performance versus energy consumption of different hardware architectures which have not been conducted as of yet in the field of hyperspectral imaging, particularly relevant to achieve processing results in real-time. They focus a thoughtful perspective on this relevant issue and analyze the performance versus energy consumption ratio of different processing chains for spectral unmixing when implemented on multi-core platforms.
- At this point, it is worthy to pay attention to an important aspect regarding low-level vision - the motion estimation - existing nowadays. It has many important applications and is crucial to the hardware acceleration of these algorithms. Two papers representing different family methods: "gradient-family" and "matching-family" are presented in this special issue.

A gradient family-based method is presented in the contribution 'Robust motion estimation on a low-power multi-core DSP' [17] authored by Francisco D. Igual et al., which addresses the efficient implementation of a robust gradient-based optical flow model in a low-power platform based on a multi-core digital signal processor (DSP C6678 DSP from Texas Instruments, Dallas, TX, USA). The aim of this work was to carry out a feasibility study on the use of these devices in autonomous systems such as robot navigation, biomedical assistance, or tracking with not only power restrictions but also real-time requirements. The interest of this research is particularly relevant in optical flow scope because this system can be considered as an alternative solution for mid-range video resolutions when a combination of in-processor parallelism with optimizations such as efficient memory-hierarchy exploitation and multi-processor parallelization are applied.

A matching family-based method is presented in the contribution 'Acceleration of block-matching algorithms using a custom instruction-based paradigm on a Nios II microprocessor' [18] authored by Diego González et al.; this paper focuses on the optimization of matching algorithms widely used for video coding standards using an Altera custom instruction-based paradigm and a combination of synchronous dynamic random access memory (SDRAM) with on-chip memory in Nios II processors. A complete profile of the algorithms is achieved before the optimization, which locates code leaks and, afterward, creates a custom instruction set, which is then added to the specific design, enhancing the original system. As well, every possible memory combination between on-chip memory and SDRAM has been tested to achieve the best performance. The final throughput of the complete designs are shown. This manuscript outlines a low-cost system, mapped using very large scale integration technology, which accelerates software algorithms by converting them into custom hardware logic blocks and showing the best combination between on-chip memory and SDRAM for the Nios II processor.

- Keeping it going at a low-level vision layer, the next contribution focuses on augmented vision for impaired people; this paper is entitled 'A reconfigurable real-time morphological system for augmented vision' [19] authored by Ryan B. Gibson et al. Augmented visual aid devices require highly user-customizable algorithm designs for subjective configuration per task, where current digital image processing visual aids offer very little user-configurable option. The high spatial frequencies of an image can be extracted by edge detection techniques and overlaid on top of the original image to improve visual perception among the visually impaired. This paper presents a highly user-reconfigurable morphological edge enhancement system on FPGA, where the morphological, internal and external edge gradients can be selected from the presented architecture with specified edge thickness and magnitude. In addition, the morphology architecture supports reconfigurable shape structuring elements and configurable morphological operations. The proposed morphology-based visual enhancement system introduces a high degree of user flexibility in addition to meeting real-time constraints capable of obtaining 93 fps for high-definition image resolution.
- Finally, the last paper contributed in this special issue is entitled 'Parallel embedded processor architecture for FPGA based image processing using parallel software skeletons' [20] authored by Hanan Chenini et al. This application belongs to the mid-

level vision, the authors provide a rapid prototyping tool as a graphic programming environment (CubeGen). Further, it offers a set of parallel software skeletons as a communication library, providing a software abstraction to enable quick implementation of complex image processing applications on the FPGA platform. The design of homogeneous network of communicating processor is presented from the hardware and software specification down to synthesizable hardware description. Additionally, the present approach is extended to support more complex applications by implementing a soft multiprocessor and show the impact of various configuration choices to match the specific application needs. Using images of a real road scene, the performance results of the road recognition algorithm on a Xilinx Virtex-6 FPGA platform not only achieve the desired latency but further improve the tracking performance which depends mainly on the number of hypotheses.

Conclusions

This editorial paper has covered the accepted contributions as well as the topics covered for the special issue 'Hardware Implementation of Machine Vision Systems: Image and Video Processing' in the *Journal on Advances Signal Processing (JASP)*, 2013.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The guest editors would like to thank the authors for their cooperation, interest, and their top-quality contributions presented in this special issue. Also, the guest editor wants to thank our Editor-in-Chief Professor Geert Leus (as well as Professor Phillip Regalia, the former Editor-in-Chief when this special issue was launched). Finally, we thank the entire Editorial Office Staff for their dedication and effort devoted to this special issue. Also, we want to acknowledge the Spanish Project TIN 2012/32180.

Author details

¹Department of Computer Architecture and Automation, Complutense University, Ciudad Universitaria s/n, 28040 Madrid, Spain. ²Department of Electrical and Computer Engineering, FAMU-FSU College of Engineering, 2525 Pottsdamer Street, 32310 Tallahassee, FL, USA.

Received: 6 September 2013 Accepted: 19 September 2013

Published: 27 September 2013

References

1. F Ayuso, G Botella, C García, M Prieto, F Tirado, GPU-based acceleration of bioinspired motion estimation model, in *Concurrency and Computation: Practice and Experience*, vol. 25 (Wiley, New York, 2013), pp. 1037–1056
2. G Botella, A García, M Rodríguez-Alvarez, E Ros, U Meyer-Bäse, MC Molina, Robust bioinspired architecture for optical-flow computation. *IEEE Trans. VLSI Syst.* **18**(4), 616–629 (2010)
3. C García, G Botella, F Ayuso, M Prieto, M Tirado, Multi-GPU based on multicriteria optimization for motion estimation system. *EURASIP J. Adv. Signal Processing* **2013**, 23 (2013)

4. G Botella, U Meyer-Bäse, A García, Bio-inspired robust optical flow processor system for VLSI implementation. *Electronics Letters* **45**(25), 1304–1305 (2009). DOI: 10.1049/el.2009.1718
5. G Botella, E Ros, M Rodríguez, A García, S Romero, Pre-processor for bioinspired optical flow models: a customizable hardware implementation, in *Proceedings of the 13th IEEE Mediterranean Electrotechnical Conference, MELECON 2006, Benalmádena, 16-18 May* (Benalmádena, 2006), pp. 93–96
6. D González, G Botella, U Meyer-Bäse, C García, C Sanz, M Prieto-Matias, F Tirado, A Low cost matching motion estimation sensor based on the NIOS II microprocessor. *Sensors* **12**, 13126–13149 (2012)
7. R Szelinsky, *Computer Vision Algorithms and Applications* (Springer, Berlin, Germany, 2011)
8. P Dudek, A Lopich, V Gruev, A pixel-parallel cellular processor array in a stacked three-layer 3D silicon-on-insulator technology, in *European Conference on Circuit Theory and Design, 2009 (ECCTD 2009, Antalya, Aug 2009)*, pp. 23–27
9. E Alcocer, O López-Granado, R Gutierrez, MP Malumbres, MPCM: a hardware coder for super slow motion video sequences. *EURASIP J. Adv Signal Processing* **2013**, 142 (2013)
10. Martínez-Rach, O Miguel, O López-Granado, V Galiano, H Migallón, J Llor, MP Malumbres, Enhancing LTW image encoder with perceptual coding and GPU-optimized 2D-DWT transform. *EURASIP J. Adv Signal Processing* **2013**, 141 (2013)
11. V Galiano, O López-Granado, LA Drummond, MP Malumbres, H Migallón, GPU-based 3D lower tree wavelet video encoder. *EURASIP J. Adv Signal Processing* **2013**, 141 (2013)
12. R Rodríguez-Sánchez, J Martínez, J Cock, G Fernández-Escribano, B Pieters, JL Sánchez, JM Claver, R De Walle, Adapting hierarchical bidirectional inter prediction on a GPU-based platform for 2D and 3D H.264 video coding. *EURASIP J. Adv Signal Processing* **2013**, 67 (2013)
13. LI Timchenko, NI Kokryatskaya, VV Shpakovych, Modeling of a method of parallel hierarchical transformation for fast recognition of dynamics images. *EURASIP J. Adv Signal Processing* **2013**, 87 (2013)
14. M Kumm, D Fanghänel, K Möller, P Zipf, U Meyer-Bäse, FIR Filter optimization for video processing on FPGAs. *EURASIP J. Adv Signal Processing* **2013**, 111 (2013)
15. B Leung, S Memik, Exploring super-resolution implementations across multiple platforms. *EURASIP J. Adv Signal Processing* **2013**, 116 (2013)
16. A Remón, S Sánchez, S Bernabé, ES Quintana-Ortí, A Plaza, Performance versus energy consumption of hyperspectral unmixing algorithms on multi-core platforms. *EURASIP J. Adv Signal Processing* **2013**, 68 (2013)
17. FD Igual, G Botella, C García, M Prieto, F Tirado, Robust motion estimation on a low-power multi-core DSP. *EURASIP J. Adv Signal Processing* **2013**, 99 (2013)
18. D González, G Botella, C García, M Prieto, F Tirado, Acceleration of block-matching algorithms using a custom instruction-based paradigm on a Nios II microprocessor. *EURASIP J. Adv Signal Processing* **2013**, 118 (2013)
19. RM Gibson, A Ahmadinia, SG McMeekin, NC Strang, G Morison, A reconfigurable real-time morphological system for argumented vision. *EURASIP J. Adv Signal Processing* **2013**, 134 (2013)
20. H Chenini, JP Derutin, R Aufrere, R Chapuis, Parallel embedded processor architecture for FPGA based image processing using parallel software Skeletons. *EURASIP J. Adv Signal Processing* (2013), in press

doi:10.1186/1687-6180-2013-152

Cite this article as: Botella et al.: Hardware implementation of machine vision systems: image and video processing. *EURASIP Journal on Advances in Signal Processing* 2013 **2013**:152.